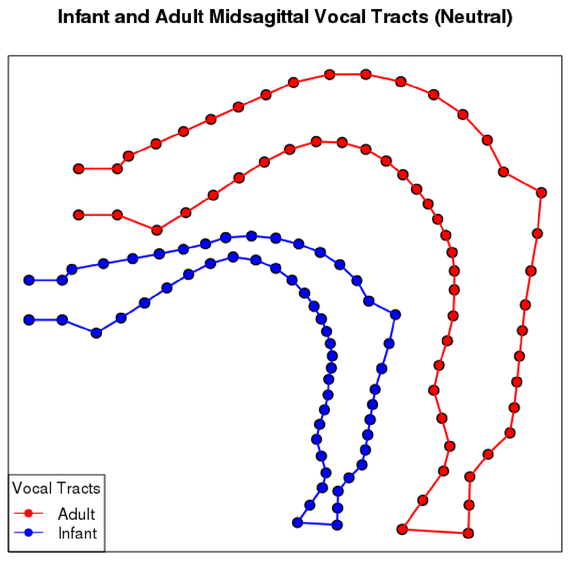
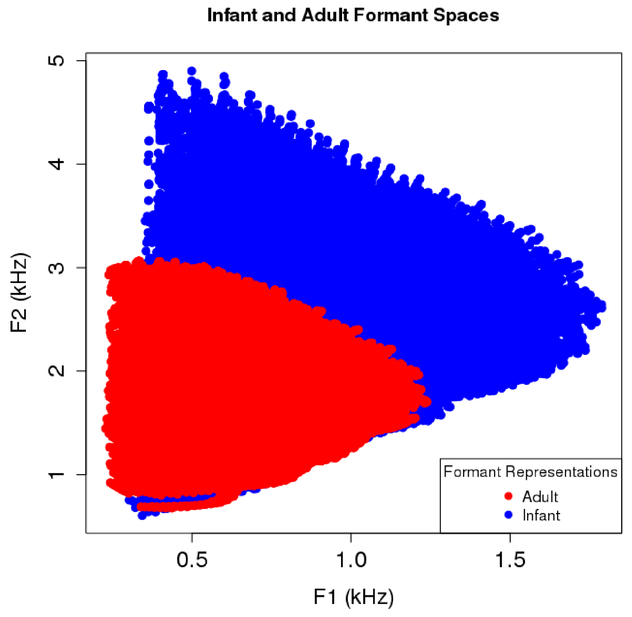
# Data and Figures

## Articulatory and Acoustic Data

- We use Maeda's and Boe's (1997) Variable Linear Articulatory Model (VLAM) to model the vowel productions of an infant and an adult caretaker.
- We use the 6 month-old setting of the VLAM as our model infant, and the 10 year-old setting for the adult caretaker (as it was perceived to be most similar to a young female adult in a cross-language perception study (Munson et al., 2010)).




## Representations

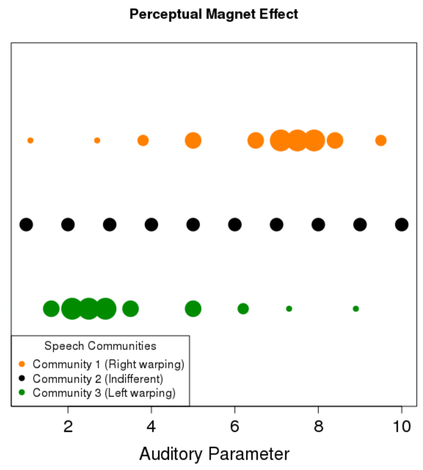
- The articulatory representations (1, left) are principal component vector arguments for the VLAM, each yielding a vowel signal with a formant representation (1, middle).

$$\mathbf{a} = \langle a_1, \ldots, a_7 \rangle, \quad \mathbf{f} = \langle f_1, f_2, f_3 \rangle, \quad \mathbf{e} = \langle e_1, \ldots, e_{361} \rangle \qquad (1)$$

- The auditory representations (1, right) are "excitation patterns" derived from the vowel signals using the transformations described in Moore et al. (1997)

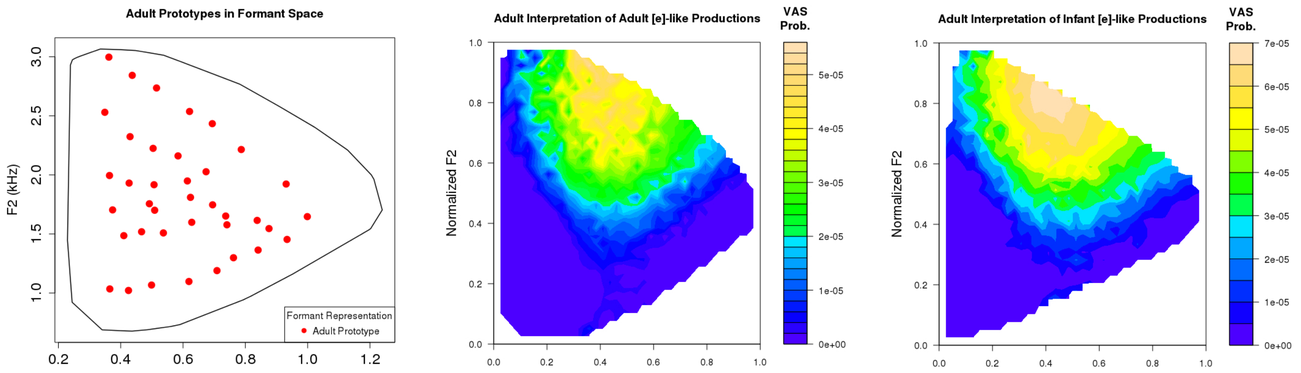
## Perceptual Magnet Effect

- The perceptual magnet effect (Kuhl, 1991) names the phenomenon wherein the perception of a vowel in a given language is influenced by "perceptual magnets" located in a perceptual metric space over representations of vowels in that language.

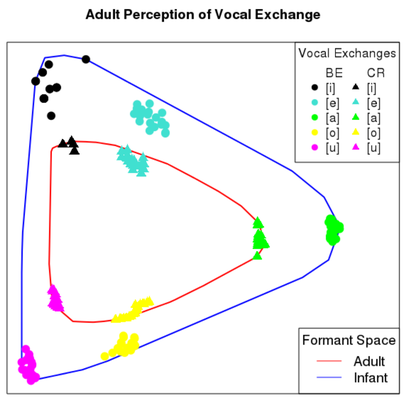


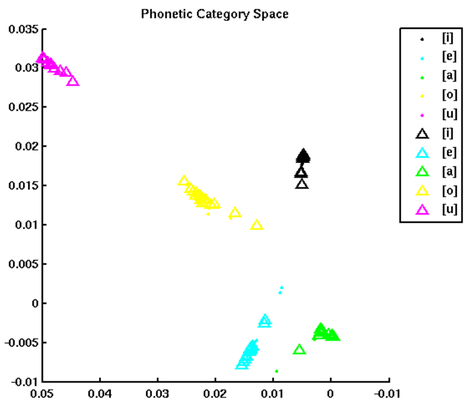

## Adult Perception of Infant Vocalizations, and Contingent Responses

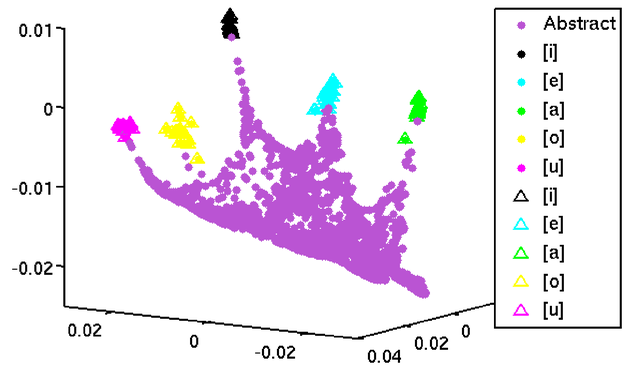
- We use 38 perceptually categorized vowel stimuli (Munson et al. 2010) generated by the VLAM set at 10 years of age to model the vowel categories of an "average" Greek-speaking caretaker.
- We additively interpolate the categorizations over the formant space of the caretaker and "project" the categories to the infant's acoustic space, modeling the caretaker's interpretations of infant vocalizations.




## Simulation Results




- In response to an infant's vocalization, the caretaker interprets and responds, yielding vocal imitation pairs (left) that guide manifold alignment.
- Cross-modal manifold alignment yields a basis for vowel categorization (middle), while the intra-modal alignment (right) yields perceptual warping.

Mail: plummer@ling.ohio-state.edu WWW: http://www.learningtotalk.org